(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 September 2006 (28.09.2006)

PCT

(10) International Publication Number
WO 2006/099719 A1

(51) International Patent Classification:
*H04L 29/02* (2006.01) *G06F 12/00* (2006.01)

(21) International Application Number:
PCT/CA2006/000363

(22) International Filing Date: 17 March 2006 (17.03.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/088,948 25 March 2005 (25.03.2005) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 11/088,948 (CON)
Filed on 25 March 2005 (25.03.2005)

(71) Applicant *(for all designated States except US)*: **ELLIPTIC SEMICONDUCTOR INC.** [CA/CA]; 362 Terry Fox Drive, Suite 220, Kanata, Ontario K2K 2P5 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BOWLER, Michael** [CA/CA]; 247 Daniel Avenue, Ottawa, Ontario K1Y 0C7 (CA). **HAMILTON, Neil** [CA/CA]; 40 Bannock Crescent, Kanata, Ontario K2K 2P9 (CA). **BORZA, Michael** [CA/CA]; 2036 Woodcrest Road, Ottawa, Ontario K1H 6H8 (CA).

(74) Agent: **FREEDMAN, Gordon**; Freedman & Associates, 117 Centrepointe Drive, Suite 350, Nepean, Ontario K2G 5X3 (CA).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, YU, ZA, ZM, ZW.

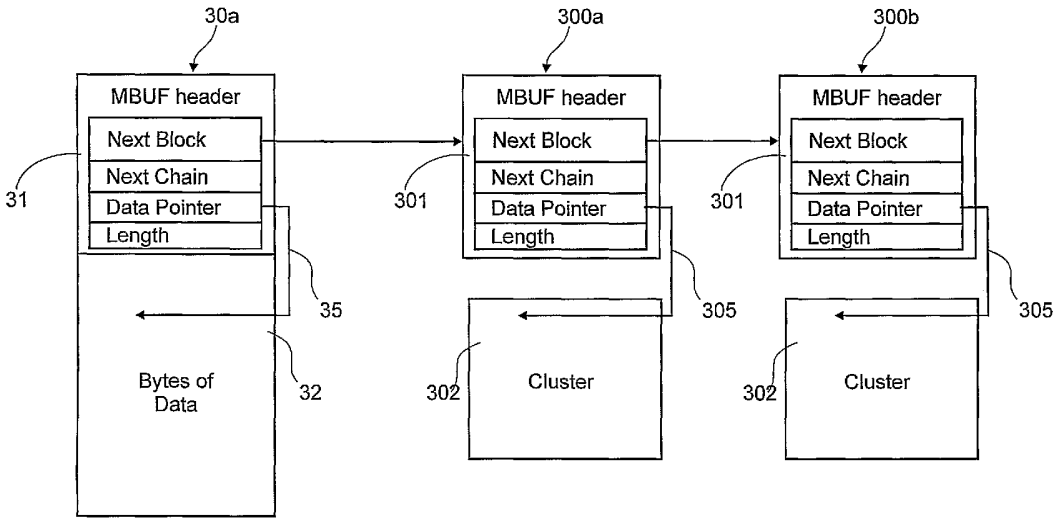
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declarations under Rule 4.17:**
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**
— *with international search report*

*[Continued on next page]*

(54) Title: PACKET MEMORY PROCESSING SYSTEM HAVING MEMORY BUFFERS WITH DIFFERENT ARCHITECTURES AND METHOD THEREFOR




(57) Abstract: An architecture for use in packet processing and supporting compatibility with current BSD implementations for packet processing is proposed wherein two MBUF formats are supported. A first format includes a header portion and a data portion for storing data therein. A second format includes a header portion but is absent a data portion and is for addressing data stored within a cluster and external to the MBUF itself.

WO 2006/099719 A1

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

ABSTRACT

An architecture for use in packet processing and supporting compatibility with current BSD implementations for packet processing is proposed wherein two MBUF formats are supported. A first format includes a header portion and a data portion for storing data therein. A second format includes a header portion but is absent a data portion and is for addressing data stored within a cluster and external to the MBUF itself.

# PACKET MEMORY PROCESSING SYSTEM HAVING MEMORY BUFFERS WITH DIFFERENT ARCHITECTURES AND METHOD THEREFOR

## FIELD OF THE INVENTION

**[001]** The invention relates to the field of data networking protocol stack processors and more specifically to the field of managing of data packets in an efficient manner.

## BACKGROUND OF THE INVENTION

**[002]** Memory buffers (MBUFs) are well known in the software and hardware design of protocol stack processors. MBUFs according to the Berkeley Software Distribution (BSD) implementation include a control portion and a data portion. The MBUF allocated memory is of a fixed size relating to the memory available and communication data being processed. A typical choice for MBUF size is between 128 and 512 bytes.

**[003]** BSD was designed with systems having significant power and memory resources in mind. In today's technology markets, wireless battery operated circuits are becoming ever more popular. With their popularity, increased resource efficiency becomes a critical limitation on performance.

**[004]** It would be advantageous to provide an architecture for more efficient resource utilization that is backward compatible with current protocol implementations.

## SUMMARY OF THE INVENTION

**[005]** In accordance with the invention there is provided a memory buffer architecture for use with a BSD implementation comprising: a first MBUF format including a control portion and a data portion, the first MBUF format for use where data is to be stored within the MBUF; and, a second MBUF format including a control portion and absent a substantial data portion, the second MBUF format for use in indexing data within a cluster.

**[006]** In accordance with another aspect of the invention there is provided a storage device comprising data stored therein, the data for resulting in implementation of an integrated circuit design including circuitry for allocating a first MBUF format including a control portion and a data portion, the first MBUF format for use where data is to be stored within the MBUF; and, a second MBUF format including a control portion and absent a data portion, the second MBUF format for use in indexing data within a cluster, the MBUFs including pointer data memory locations within a control portion thereof for being arranged in a linked list.

**[007]** In accordance with another aspect of the invention there is provided a method of packet processing comprising: storing packet data within a linked list, the linked list including MBUFs linked together at least one of the linked MBUFs addressing data within a cluster external to said MBUF, wherein at least one of the at least one of the linked MBUFs comprises a control portion absent a data portion thereof, and wherein at least one of the other than the at least one of the linked MBUFs comprises a control portion and a data portion.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[008]** Exemplary embodiments of the invention will now be described in conjunction with the following drawings, in which:

**[009]** Fig. 1 illustrates a plurality of prior art memory buffers (MBUFs);

**[0010]** Fig. 2 illustrates a plurality of prior art MBUFs in a chain and addressing data within a cluster;

**[0011]** Fig. 3 illustrates two MBUFs having different formats in accordance with the present invention; and,

**[0012]** Fig. 4 illustrates a chain of MBUFs addressing data within a cluster, some of the MBUFs having a first form and some of the MBUFs having a second other form.

## DETAILED DESCRIPTION THE INVENTION

**[0013]** Fig. 1 illustrates a chain consisting of two MBUFs 10 according to the prior art. Each MBUF comprises control data 11 in the form of an MBUF header comprising a link address for a next block 13, a link address for a next chain 14, a data address 15, and a data length 16. Typically, the MBUF header occupies 32 bytes. Each MBUF further comprises a data portion 12 for storing of data therein. Typically, MBUFs are between 128 and 512 bytes with an MBUF of 256 bytes being most typical. As such, the use of MBUFs is relatively memory efficient for small packets.

**[0014]** Referring to Fig. 2, for larger data packets, a cluster 21 is often used wherein the data portion 22 of the MBUF 10 is unused and the cluster 21 – a separate data location within memory - is used for storing of the data. The control portion 11 of the MBUF 10 remains similarly utilized though the data pointer 15 points to an address within the cluster. In this fashion, the memory usage is not overly represented by the control data portion 11 of the MBUF 10 and the null MBUF data portion 22 is small compared to the amount of data stored within the cluster 21. Thus, memory usage efficiency remains regardless of packet data size.

**[0015]** The prior art MBUF structure suffers from considerable drawbacks for very high efficiency operation. Firstly, the memory storage wasted by unused data memory both in MBUFs relying on data within clusters and within MBUFs whose data portions are other than completely utilized is of concern. Also, allocating more memory than necessary can be problematic in integrated devices where available memory is highly constrained. Thirdly, in cases where MBUF reorganization is necessary, efficiency is further reduced.

**[0016]** Referring to Fig. 3, an MBUF architecture is shown including two distinct MBUF formats 30 and 300. The first MBUF format 30 is a typical MBUF format with a control portion 31 and a data portion 32 having a memory size selected for memory utilization efficiency. A pointer 35 from the control portion 31 addresses the data within the data portion 32. The second MBUF format 300 includes only the control portion 301 for use in addressing data stored within a cluster 302 and typically occupies only 32 bytes.

**[0017]** The enhanced memory efficiency of the two formats of MBUFs 30 and 300 is highly advantageous. For example, the format is compatible with existing protocols and BSD software since it does not allocate an unused resource. Modifying existing implementations in a fashion that is backwards compatible is often desirable.

**[0018]** Referring to Fig. 4, a chain of MBUFs is shown comprising both types of MBUFs in accordance with the invention. A first MBUF in the chain 30a, is an MBUF of the first type including both an MBUF control portion 31 and a data portion 32. Pointer 35 within the MBUF 30a addresses data within the data portion 32 of the MBUF 30a. A next MBUF pointer within the MBUF control portion 31 addresses a subsequent MBUF 300a. The subsequent MBUF 300a consists of an MBUF control portion 301. Alternatively, the MBUF 300a comprises an MBUF control portion and is absent a data portion for storing of data therein. The data pointer 305 within the MBUF 300a addresses data within cluster 302. A next MBUF pointer within the MBUF control portion 301 addresses a subsequent MBUF 300b within the chain of MBUFs. The subsequent MBUF 300b consists of an MBUF control portion 301. Alternatively, the MBUF 300b comprises an MBUF control portion and is absent a data portion for storing of data therein. The data pointer within the MBUF 300b addresses data within a cluster 302. The next block pointer and the next chain pointer within MBUF 300b are each indicative of a lack of further MFUFs within the chain and of a lack of further chains linked to the MBUF chain, respectively.

**[0019]** Numerous other embodiments may be envisaged without departing from the spirit or scope of the invention.

# CLAIMS

What is claimed is:

1. A memory buffer architecture for use in accordance with a BSD networking implementation comprising:

a first MBUF format including a control portion and a data portion, the first MBUF format for use where data is to be stored within the MBUF; and,

a second MBUF format including a control portion and absent a substantial data portion, the second MBUF format for use in indexing data within a cluster.

2. A memory buffer architecture according to claim 1, wherein the second MBUF format is absent a data portion thereof.

3. A memory buffer architecture according to claim 2, wherein the second MBUF format is for use in indexing data within clusters and wherein the first MBUF format is for storing of data within the data portion thereof.

4. A memory buffer architecture according to claim 3, wherein in use the MBUFs are allocated and arranged as a linked list.

5. A memory buffer architecture according to claim 4, wherein the linked list is a two dimensional linked list.

6. A memory buffer architecture according to claim 4, wherein the linked list of MBUFs is representative of a single communication packet.

7. A memory buffer architecture according to claim 4, wherein the linked list of MBUFs is representative of a single communication stream.

8. A storage device comprising data stored therein, the data for resulting in implementation of an integrated circuit design including circuitry for allocating

a first MBUF format including a control portion and a data portion, the first MBUF format for use where data is to be stored within the MBUF; and,

a second MBUF format including a control portion and absent a substantial data portion, the second MBUF format for use in indexing data within a cluster,

the MBUFs including pointer data memory locations within a control portion thereof for being arranged in a linked list.

9. A storage device according to claim 8, wherein the integrated circuit is for use in packet processing in accordance with a BSD implementation, the MBUFs forming MBUFs operational in accordance with said BSD implementation.

10. A storage device according to claim 8, having further data stored therein, the data for resulting in implementation of an integrated circuit design including circuitry for allocating one of the first and second MBUF formats in dependence upon received data wherein one of the first and second MBUF formats is selected for being associated with data within the received data, said one being determined upon receipt of the data.

11. A storage device according to claim 8, having further data stored therein, the data for resulting in implementation of an integrated circuit design including circuitry for allocating one of the first and second MBUF formats in dependence upon received data wherein one of the first and second MBUF formats is selected for being associated with data within the received data, said one being determined upon receipt of the data, the second MBUF format selected for indexing data within at least a cluster.

12. A storage device according to claim 8, having further data stored therein, the data for resulting in implementation of an integrated circuit design including circuitry for allocating one of the first and second MBUF formats in dependence upon received data wherein one of the first and second MBUF formats is selected for being associated with data within the received data, said one being determined upon receipt of the data, the second MBUF format selected for indexing data within at least a cluster and the first MBUF format selected for at least one of pre-allocating memory data storage for later use and for storing of data within a data portion thereof..

13. A method of packet processing comprising:

storing packet data within a linked list, the linked list including MBUFs linked together an MBUF of the linked MBUFs addressing data within a cluster external to said MBUF, wherein said MBUF comprises a control portion absent a substantial data portion thereof, and wherein at least another MBUF of the linked MBUFs comprises a control portion and a data portion.

14. A method according to claim 13, wherein said MBUF is absent a data portion.

15. A method according to claim 14, comprising:

receiving data; and,

allocating an MBUF format selected from a first MBUF having a control portion and data portion and a second MBUF having a control portion and absent a substantial data portion in dependence upon the received data wherein the second MBUF is selected for indexing data within clusters.

16. A method according to claim 14, comprising:

receiving data; and,

allocating an MBUF format selected from a first MBUF having a control portion and data portion and a second MBUF having a control portion and absent a substantial data portion in dependence upon the received data wherein the second MBUF is selected for indexing data within clusters, said MBUF format selected upon receipt of the data.

17. A method according to claim 13, wherein the method is for use in packet processing compatible with the BSD implementation, each MBUF within the linked MBUFs forming MBUFs operational in accordance with said implementation.

18. A method according to claim 13, comprising:

receiving data; and,

allocating an MBUF format selected from a first MBUF having a control portion and data portion and a second MBUF having a control portion and absent a substantial data portion in dependence upon the received data wherein the second MBUF is selected for indexing data within clusters.

19. A method according to claim 13, comprising:

receiving data; and,

allocating an MBUF format selected from a first MBUF having a control portion and data portion and a second MBUF having a control portion and absent a substantial data portion in dependence upon the received data wherein the second MBUF is selected for indexing data within clusters, said MBUF format selected upon receipt of the data.

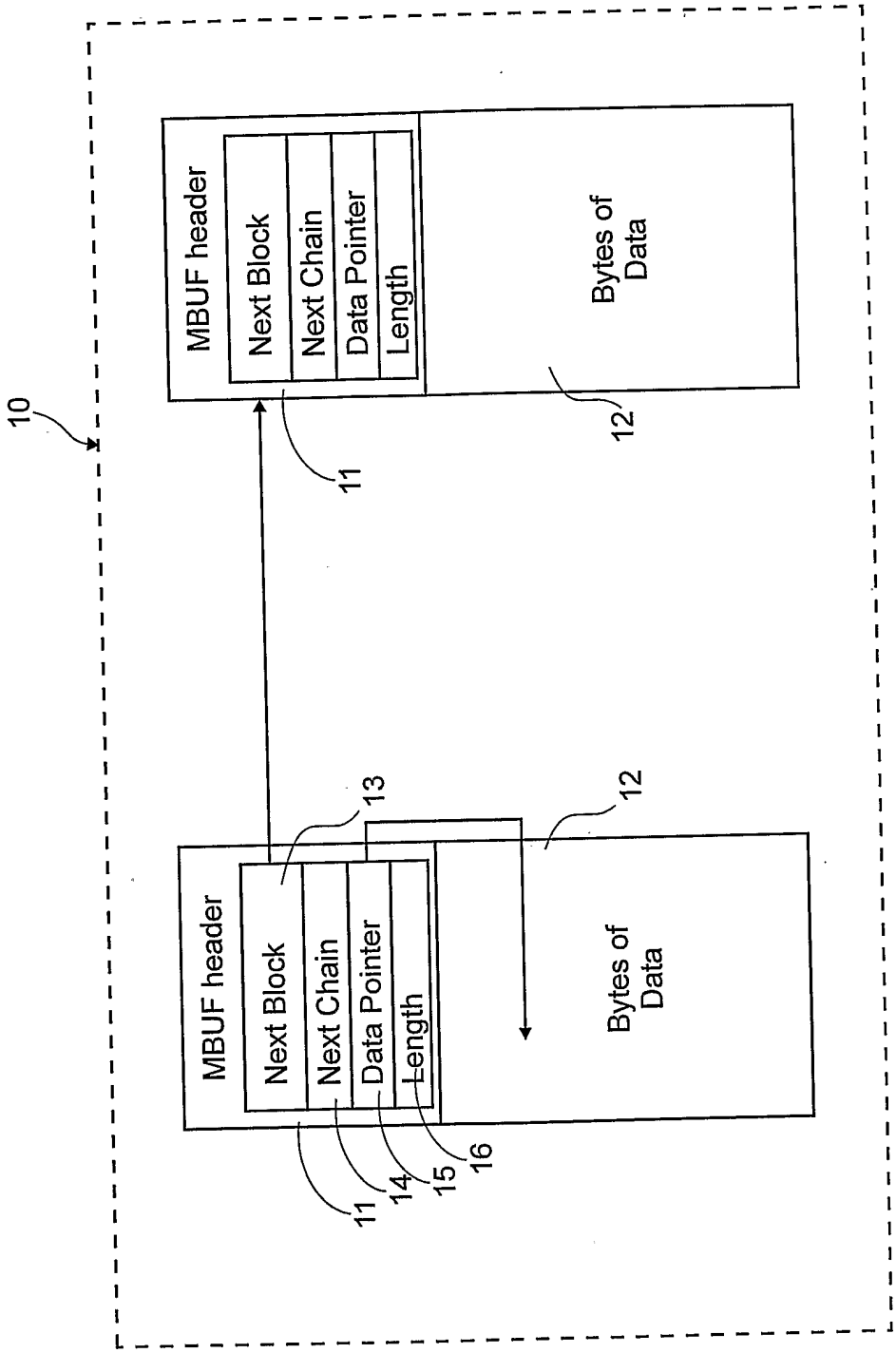



Figure 1
(PRIOR ART)

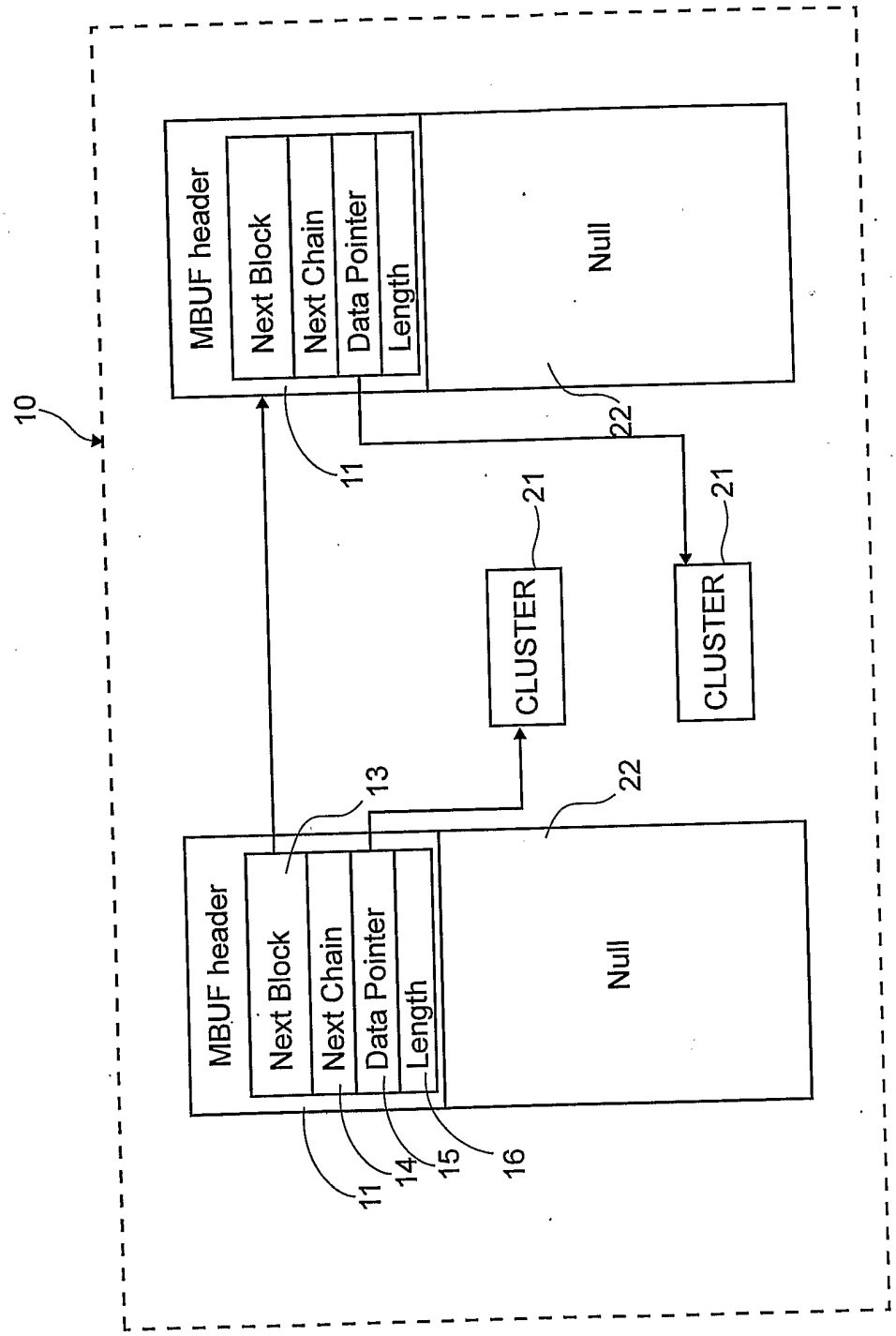
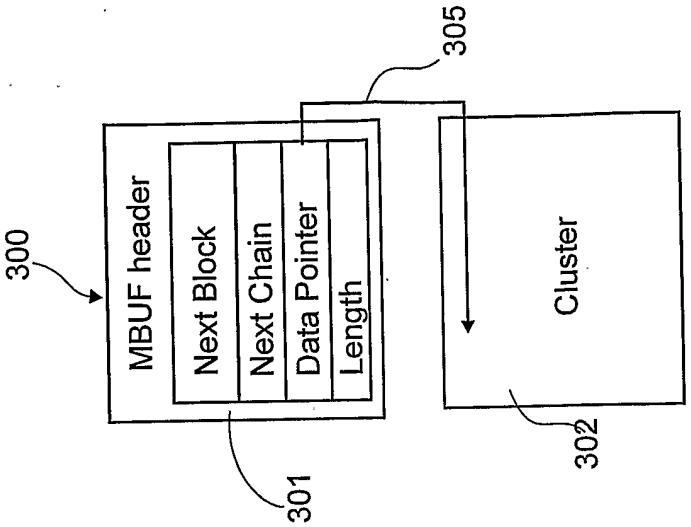



Figure 2
(PRIOR ART)


300
MBUF header
Next Block
Next Chain
Data Pointer
Length
305
Cluster
301
302

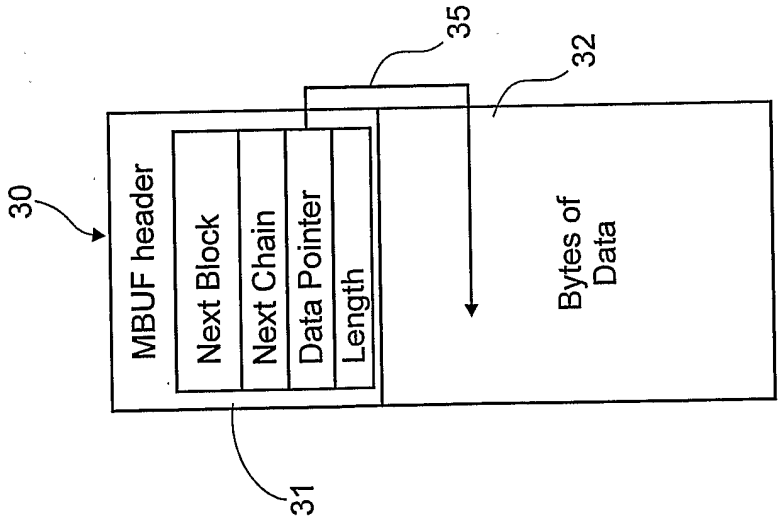

30
MBUF header
Next Block
Next Chain
Data Pointer
Length
35
32
Bytes of Data
31

Figure 3

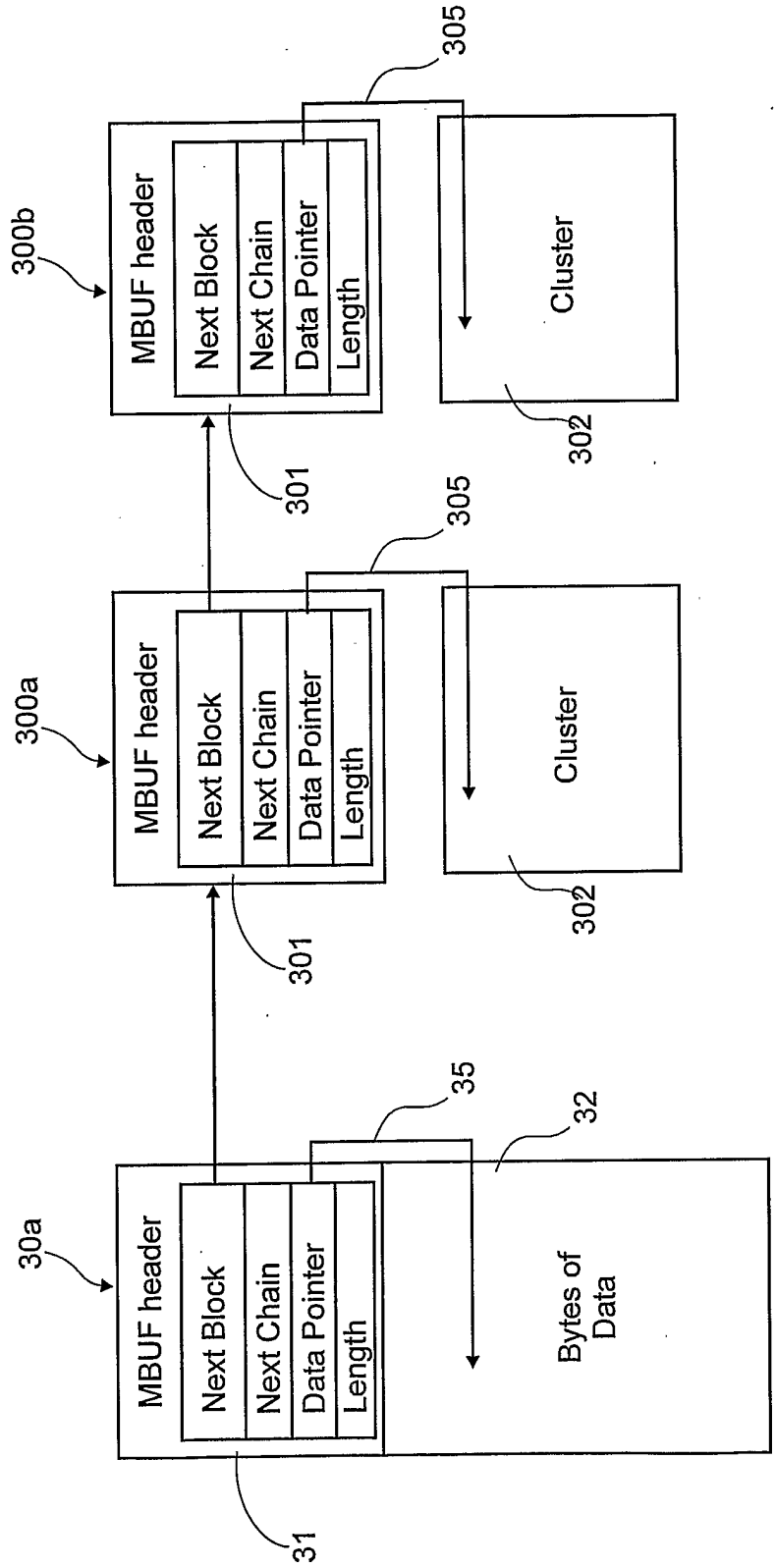

300b
MBUF header
Next Block
Next Chain
Data Pointer
Length
305
Cluster
302
301
300a
MBUF header
Next Block
Next Chain
Data Pointer
Length
305
Cluster
302
301
30a
MBUF header
Next Block
Next Chain
Data Pointer
Length
35
32
Bytes of Data
31

Figure 4

**INTERNATIONAL SEARCH REPORT**

International application No.
PCT/CA2006/000363

A. CLASSIFICATION OF SUBJECT MATTER
IPC: ***H04L 29/02*** (2006.01) , ***G06F 12/00*** (2006.01)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC: H04L, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
Google, Delphion: mbuf, memory buffer, architecture, control, data, header, format, structure, bsd

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | Leffler et al., "Networking Implementation Notes 4.4BSD Edition", computer Systems Research Group, University of California, Berkeley, June 10, 1993 (1993.06.10). Downloaded from http://docs.freebsd.org/44doc/smm/18.net/paper.pdf on July 10, 2006. | 1, 8-13, 17-19 |
| Y | | 2-7, 14-16 |
| Y | US 2003/0108045 A1, Jayam et al., June 12, 2003 (2003.06.12), abstract, paragraphs [0023-0024], [0035-0036]. | 2-7, 14-16 |
| Y | US 6,412,045 B1, DeKoning et al., June 25, 2002 (2002.06.25), whole document | 2-7, 14-16 |

[ ] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 26 July 2006 (26-07-2006) | 28 July 2006 (28-07-2006) |

| Name and mailing address of the ISA/CA | Authorized officer |
|---|---|
| Canadian Intellectual Property Office<br>Place du Portage I, C114 - 1st Floor, Box PCT<br>50 Victoria Street<br>Gatineau, Quebec K1A 0C9<br>Facsimile No.: 001(819)953-2476 | Stuart Ginn (819) 934-5147 |

| Patent Document Cited in Search Report | Publication Date | Patent Family Member(s) | Publication Date |
|---|---|---|---|
| US2003108045 | 12-06-2003 | CN1432815 A | 30-07-2003 |
| | | EP1421494 A1 | 26-05-2004 |
| | | EP1421500 A1 | 26-05-2004 |
| | | JP2005502125T T | 20-01-2005 |
| | | JP2005503699T T | 03-02-2005 |
| | | US6760769 B2 | 06-07-2004 |
| | | US6981014 B2 | 27-12-2005 |
| | | US2003061505 A1 | 27-03-2003 |
| | | US2003115338 A1 | 19-06-2003 |
| | | US2004187691 A1 | 30-09-2004 |
| | | WO03021443 A1 | 13-03-2003 |
| | | WO03021447 A1 | 13-03-2003 |
| | | WO03021452 A1 | 13-03-2003 |
| US6412045 | 25-06-2002 | DE69616148D D1 | 29-11-2001 |
| | | DE69616148T T2 | 07-03-2002 |
| | | EP0744696 A1 | 27-11-1996 |
| | | EP0744697 A1 | 27-11-1996 |
| | | JP8314848 A | 29-11-1996 |
| | | JP8339329 A | 24-12-1996 |
| | | US5917723 A | 29-06-1999 |